

**AMENDMENTS TO THE CLAIMS**

**This listing of claims will replace all prior versions and listings of claims in the application:**

**LISTING OF CLAIMS:**

Claims 1-6 (canceled)

7. (currently amended) ~~The questionnaire analysis system according to claim 4, A questionnaire analysis system comprising:~~

means for inputting a questionnaire statement including free reply description in natural language;

a network for transmitting a questionnaire reply statement,

a database for accumulating said questionnaire reply statements transmitted through said network; and

a text classification engine for reading out said questionnaire reply statements from said database and for learning a rule for classifying said questionnaire reply statement,

wherein said text classification engine includes:

morpheme analysis means for analyzing morphemes in all sentences in said questionnaire reply statement accumulated in said database;

category-text designating means for designating a category and text;

attribute selecting means for selecting attributes in plural questionnaire reply statements

being read out from said database;

rule learning means for learning said rule for expressing said correspondence of text and  
category on the basis of said words selected by attributes by said attribute selecting means; and

rule output means for issuing said rule learned by said rule learning means,

wherein said attribute selecting means computes a difference  $\Delta SC(\omega)$  between a stochastic complexity (SC) of a ~~test~~text set without consideration of appearance of a word and a stochastic complexity (SC) of a text set with consideration thereof, in each word  $\omega$  appearing in said text, and then selects said difference  $\Delta SC(\omega)$  as an attribute when said difference  $\Delta SC(\omega)$  is ~~lager~~larger than ~~said~~a threshold  $\tau$ .

8. (currently amended) ~~The questionnaire analysis system according to claim 5, A~~  
questionnaire analysis system comprising:

means for inputting a questionnaire statement including free reply description in natural  
language;

a network for transmitting a questionnaire reply statement,

a database for accumulating said questionnaire reply statements transmitted through said  
network; and

a text classification engine for reading out said questionnaire reply statements from said database and for learning a rule for classifying said questionnaire reply statement,

wherein said text classification engine includes:

morpheme analysis means for analyzing morphemes in all sentences in said questionnaire reply statement accumulated in said database;

category-text designating means for designating said category and text;

attribute selecting means for selecting attributes in plural questionnaire reply statements being read out from said database;

rule learning means for learning said rule for expressing said correspondence of text and category on the basis of said words selected by attributes by said attribute selecting means; and

rule output means for issuing said rule learned by said rule learning means,

wherein said attribute selecting means computes a difference  $\Delta SC(\omega)$  between a stochastic complexity (SC) of a test-text set without consideration of appearance of a word and a stochastic complexity (SC) of a text set with consideration thereof, in each word  $\omega$  appearing in said text, and then selects said difference  $\Delta SC(\omega)$  as an attribute when said difference  $\Delta SC(\omega)$  is lager-larger than said-a threshold  $\tau$ .

9. (currently amended) The questionnaire analysis system according to claim 6, A questionnaire analysis system comprising:

means for inputting a questionnaire statement including free reply description in natural language;

a network for transmitting said questionnaire reply statement;

a database for accumulating said questionnaire reply statement transmitted through said network;

a text classification engine for reading out said questionnaire reply statement from said database and for learning a rule for classifying said questionnaire reply statement; and

means for distributing said rule through said network according to a request from a claimant,

wherein said text classification engine includes:

morpheme analysis means for analyzing morphemes in all sentences in said questionnaire reply statement accumulated in said database;

category-text designating means for designating said category and text;

attribute selecting means for selecting attributes in plural questionnaire reply statements being read out from said database;

rule learning means for learning said rule for expressing said correspondence of text and category on the basis of said words selected by attributes by said attribute selecting means; and

rule output means for issuing said rule learned by said rule learning means,

wherein said attribute selecting means computes a difference  $\Delta SC(\omega)$  between a stochastic complexity (SC) of a ~~test~~text set without consideration of appearance of a word and a stochastic complexity (SC) of a text set with consideration thereof, in each word  $\omega$  appearing in said text, and then selects said difference  $\Delta SC(\omega)$  as an attribute when said difference  $\Delta SC(\omega)$  is ~~lager~~larger than ~~said~~a threshold  $\tau$ .

10. (currently amended) ~~A questionnaire analysis system according to claim 4, A questionnaire analysis system comprising:~~

means for inputting a questionnaire statement including free reply description in natural language;

a network for transmitting a questionnaire reply statement,

a database for accumulating said questionnaire reply statements transmitted through said network; and

a text classification engine for reading out said questionnaire reply statements from said database and for learning a rule for classifying said questionnaire reply statement,

wherein said text classification engine includes:

morpheme analysis means for analyzing morphemes in all sentences in said questionnaire reply statement accumulated in said database;

category-text designating means for designating a category and text;

attribute selecting means for selecting attributes in plural questionnaire reply statements being read out from said database;

rule learning means for learning said rule for expressing said correspondence of text and category on the basis of said words selected by attributes by said attribute selecting means; and

rule output means for issuing said rule learned by said rule learning means,

wherein said rule learning means:

forms said text set by replacing with an expression of  $(d_1, c_1), (d_2, c_2), \dots, (d_m, c_m)$  [where each  $d_i$  is a multi-valued discrete vector  $d_i = (\omega_{i1}, \omega_{i2}, \dots, \omega_{in})$  ( $i = 1, \dots, m$ ),  $\omega_{ij}$  is 1 when word obtained by attribute selection  $\omega_j$  ( $j = 1, \dots, n$ ) appears in said  $i$ -th text, or 0 otherwise,  $c_i$  expresses said value (label) of said category according to said  $i$ -th text and each  $c_i$  is 1 when belonging to a specific category, or 0 otherwise, and  $m$  is said number of texts];

selects said rules of if-then-else format and sequentially adds said selected rules to said stochastic decision list by employing said information quantity standard such as said extended stochastic complexity (SC) minimum principle or SC minimizing principle; and

removes said rules one by one from said last one of said stochastic decision list, and clips continuously until none should be removed from said viewpoint of said extended SC minimum principle.

11. (currently amended) ~~A questionnaire analysis system according to claim 5, A~~  
questionnaire analysis system comprising:

means for inputting a questionnaire statement including free reply description in natural  
language;

a network for transmitting a questionnaire reply statement,  
a database for accumulating said questionnaire reply statements transmitted through said  
network; and

a text classification engine for reading out said questionnaire reply statements from said  
database and for learning a rule for classifying said questionnaire reply statement,

wherein said text classification engine includes:  
morpheme analysis means for analyzing morphemes in all sentences in said questionnaire  
reply statement accumulated in said database;

category-text designating means for designating said category and text;  
attribute selecting means for selecting attributes in plural questionnaire reply statements  
being read out from said database;

rule learning means for learning said rule for expressing said correspondence of text and  
category on the basis of said words selected by attributes by said attribute selecting means; and  
rule output means for issuing said rule learned by said rule learning means,

wherein said rule learning means:

forms said text set by replacing with an expression of  $(d_1, c_1), (d_2, c_2), \dots, (d_m, c_m)$  [where each  $d_i$  is a multi-valued discrete vector  $d_i = (\omega_{i1}, \omega_{i2}, \dots, \omega_{in})$  ( $i = 1, \dots, m$ ),  $\omega_{ij}$  is 1 when word obtained by attribute selection  $\omega_j$  ( $j = 1, \dots, n$ ) appears in said  $i$ -th text, or 0 otherwise,  $c_i$  expresses said value (label) of said category according to said  $i$ -th text and each  $c_i$  is 1 when belonging to a specific category, or 0 otherwise, and  $m$  is said number of texts];

selects said rules of if-then-else format and sequentially adds said selected rules to said stochastic decision list by employing said information quantity standard such as said extended stochastic complexity (SC) minimum principle or SC minimizing principle; and

removes said rules one by one from said last one of said stochastic decision list, and clips continuously until none should be removed from said viewpoint of said extended SC minimum principle.

12. (currently amended) ~~A questionnaire analysis system according to claim 6, A~~  
questionnaire analysis system comprising:

means for inputting a questionnaire statement including free reply description in natural language;

a network for transmitting said questionnaire reply statement;

a database for accumulating said questionnaire reply statement transmitted through said network;

a text classification engine for reading out said questionnaire reply statement from said database and for learning a rule for classifying said questionnaire reply statement; and means for distributing said rule through said network according to a request from a claimant,

wherein said text classification engine includes:

morpheme analysis means for analyzing morphemes in all sentences in said questionnaire reply statement accumulated in said database;  
category-text designating means for designating said category and text;  
attribute selecting means for selecting attributes in plural questionnaire reply statements being read out from said database;

rule learning means for learning said rule for expressing said correspondence of text and category on the basis of said words selected by attributes by said attribute selecting means; and rule output means for issuing said rule learned by said rule learning means.

wherein said rule learning means:

forms said text set by replacing with an expression of  $(d_1, c_1), (d_2, c_2), \dots, (d_m, c_m)$  [where each  $d_i$  is a multi-valued discrete vector  $d_i = (\omega_{i1}, \omega_{i2}, \dots, \omega_{in})$  ( $i = 1, \dots, m$ ),  $\omega_{ij}$  is 1 when word obtained by attribute selection  $\omega_j$  ( $j = 1, \dots, n$ ) appears in said  $i$ -th text, or 0 otherwise,  $c_i$  expresses said value (label) of said category according to said  $i$ -th text and each  $c_i$  is 1 when belonging to a specific category, or 0 otherwise, and  $m$  is said number of texts];

selects said rules of if-then-else format and sequentially adds said selected rules to said stochastic decision list by employing said information quantity standard such as said extended stochastic complexity (SC) minimum principle or SC minimizing principle; and

removes said rules one by one from said last one of said stochastic decision list, and clips continuously until none should be removed from said viewpoint of said extended SC minimum principle.

13. (currently amended) A computer program product for analyzing questionnaire reply which comprises:

a morpheme analysis procedure for analyzing morphemes in all sentences in said questionnaire reply statements accumulated in a database;

a category-text designating procedure for designating said-a category and text in said-a text classification engine;

an attribute selecting procedure for selecting attributes in plural questionnaire reply statements being read out from said database;

a rule learning means for learning said rule for expressing said correspondence of text and category on said basis of said words selected by attributes by said attribute selecting procedure; and

a rule output procedure for issuing said rule learned by said rule learning procedure,

wherein said attribute selecting procedure computes a difference  $\Delta SC(\omega)$  between a stochastic complexity (SC) of a text set without consideration of appearance of a word and a stochastic complexity (SC) of a text set with consideration thereof, in each word  $\omega$  appearing in said text, and then selects said difference  $\Delta SC(\omega)$  as an attribute when said difference  $\Delta SC(\omega)$  is larger than a threshold  $\tau$ .

14. (new) A questionnaire analysis system comprising:

a database for accumulating natural language questionnaire reply statements; and  
a text classification engine that reads out said questionnaire reply statements from said database and formulates a rule for classifying said questionnaire reply statements, said text classification engine having an attribute selector that selects attributes from a plurality of questionnaire reply statements read out from said database,

wherein said attribute selector computes a difference between a stochastic complexity of a text set without consideration of an appearance of a selected word and a stochastic complexity of the text set with consideration of the selected word, and selects said difference as an attribute when said difference is larger than a threshold value.

15. (new) The questionnaire analysis system according to claim 14, said text classification engine further comprising:

a morpheme analyzer that analyzes morphemes in all sentences in said questionnaire reply statements accumulated in said database;

category-text designating means for designating a category and the text set used by said attribute selector; and

a rule formulator that formulates said rule for expressing a correspondence of the text set and the category, on the basis of the selected word of said attribute selector.

16. (new) A computer program product for analyzing questionnaire reply which comprises:

a morpheme analysis procedure for analyzing morphemes in all sentences in said questionnaire reply statements accumulated in a database;

a category-text designating procedure for designating said category and text in said text classification engine;

an attribute selecting procedure for selecting attributes in plural questionnaire reply statements being read out from said database;

a rule learning means for learning said rule for expressing said correspondence of text and category on said basis of said words selected by attributes by said attribute selecting procedure; and

a rule output procedure for issuing said rule learned by said rule learning procedure,

wherein said rule learning means:

forms said text set by replacing with an expression of  $(d_1, c_1), (d_2, c_2), \dots, (d_m, c_m)$  [where each  $d_i$  is a multi-valued discrete vector  $d_i = (\omega_{i1}, \omega_{i2}, \dots, \omega_{in})$  ( $i = 1, \dots, m$ ),  $\omega_{ij}$  is 1 when word obtained by attribute selection  $\omega_j$  ( $j = 1, \dots, n$ ) appears in said  $i$ -th text, or 0 otherwise,  $c_i$  expresses said value (label) of said category according to said  $i$ -th text and each  $c_i$  is 1 when belonging to a specific category, or 0 otherwise, and  $m$  is said number of texts];

selects said rules of if-then-else format and sequentially adds said selected rules to said stochastic decision list by employing said information quantity standard such as said extended stochastic complexity (SC) minimum principle or SC minimizing principle; and

removes said rules one by one from said last one of said stochastic decision list, and clips continuously until none should be removed from said viewpoint of said extended SC minimum principle.